

CBO TESTIMONY

**Statement of
Peter R. Orszag
Director**

Research on the Comparative Effectiveness of Medical Treatments: Options for an Expanded Federal Role

**before the
Subcommittee on Health
Committee on Ways and Means
U.S. House of Representatives**

June 12, 2007

*This document is embargoed until it is delivered at 10:00 a.m. (EDT), Tuesday, **June 12, 2007**. The contents may not be published, transmitted, or otherwise communicated by any print, broadcast, or electronic media before that time.*



**CONGRESSIONAL BUDGET OFFICE
SECOND AND D STREETS, S.W.
WASHINGTON, D.C. 20515**

Chairman Stark and Members of the Committee, it is my pleasure to be here today to discuss an issue that could play an important role in helping to address the central fiscal challenge facing the nation: rising health care costs. Over the past four decades, Medicare's and Medicaid's costs per beneficiary have increased about 2.5 percentage points faster per year than has per capita gross domestic product (GDP).¹ If those costs continued growing at the same rate over the next four decades, federal spending on those two programs alone would rise from 4.5 percent of GDP today to roughly 20 percent by 2050. The rate at which health care costs grow relative to income is the most important determinant of the long-term fiscal balance; it exerts a significantly larger influence on the budget over the long term than other commonly cited factors, such as the aging of the population.

Rising health care costs represent a challenge not only for the federal government but also for private payers. Indeed, both trends largely reflect the same underlying forces, and cost growth per beneficiary in Medicare and Medicaid has tracked that in the rest of the health system over long periods of time. Total health care spending, which consumed about 8 percent of the U.S. economy in 1975, currently accounts for about 16 percent of GDP, and that share is projected to reach nearly 20 percent by 2016. About half of that spending is now publicly financed, and half is privately financed.

A variety of evidence suggests opportunities to constrain health care costs both in the public programs and in the rest of the health system without adverse health consequences. Perhaps the most compelling evidence of those opportunities involves the substantial geographic differences in spending on health care—both among countries and within the United States—which do not translate into higher life expectancy or measured improvements in other health statistics in the higher-spending regions. For example, Medicare's costs per beneficiary vary significantly in different regions of the United States. Research has shown that much of the variation cannot be explained by differences in the population and that the higher-spending regions do not generate better health outcomes than the lower-spending regions.²

Furthermore, hard evidence is often unavailable about which treatments work best for which patients or whether the added benefits of more-effective but more-expensive services are sufficient to warrant their added costs. In many cases, the

1. That figure takes changes in the age and sex of enrollees into account for Medicare but not for Medicaid. See Congressional Budget Office, *The Long-Term Budget Outlook* (December 2005), pp. 6–7 and 31–32.

2. See John E. Wennberg, Elliot S. Fisher, and Jonathan S. Skinner, "Geography and the Debate Over Medicare Reform," *Health Affairs*, Web Exclusive (February 13, 2002), pp. w96–w114; and Elliot S. Fisher and others, "The Implications of Regional Variations in Medicare Spending, Part 1: The Content, Quality, and Accessibility of Care," *Annals of Internal Medicine*, vol. 38, no. 4 (February 18, 2003), pp. 273–287.

extent of the variation in treatments is greatest for those types of care for which evidence about relative effectiveness is lacking. Together, those findings suggest that better information about the costs and benefits of different treatment options, combined with new incentive structures reflecting the information, could eventually yield lower health care spending without having adverse effects on health—and that the potential reduction in spending below projected levels could be substantial. Moving the nation toward that possibility—which will inevitably be an iterative process in which policy steps are tried, evaluated, and reconsidered—is essential to putting the country on a sounder long-term fiscal path. But even if it did not bring about significant reductions in spending, more information about comparative effectiveness could yield better health outcomes from the resources devoted to health care.

In response to a request from the Senate Budget and Finance Committees, the Congressional Budget Office (CBO) will issue a report on comparative effectiveness in the near future. Although that report will analyze the issues surrounding federal research efforts in greater depth, CBO's preliminary work has identified several key questions. In that light, my testimony today makes five main points:

- Because any private-sector entity (such as a health plan) has only a limited incentive to produce information that could benefit many entities, an argument can be made to coordinate comparative effectiveness research in a more systematic way than is currently done. In addition, because federal health insurance programs play a large role in financing medical care and represent a significant expenditure, the federal government itself has an interest in evaluations of the effectiveness of different health care approaches.
- If policymakers want to expand federal efforts to study comparative effectiveness, the effort could be organized in different ways—for instance, by augmenting an existing agency, by establishing a new agency, by supporting an existing quasi-governmental organization, or by creating a new public-private partnership. The choice of organizational arrangement—as well as the mechanism used to provide any federal funds to it—would affect both the entity's independence and its accountability.
- The level of funding required for a new or augmented entity would depend largely on what its additional activities would involve. Synthesizing existing studies or analyzing available data on medical claims would be less expensive than conducting new head-to-head clinical trials to compare treatments but could also yield less definitive results. Having more health records available in electronic form would facilitate the use of existing data for research (if privacy concerns could also be addressed).

- To affect medical treatment and reduce health care spending in any meaningful way, the results of comparative effectiveness analyses would have to be used in ways that changed the behavior of doctors, other health professionals, and patients. For example, the higher-value care identified by comparative effectiveness research could be promoted in the health system through financial incentives—the payments doctors receive or the cost sharing that patients face. Making substantial changes in payment policies or coverage under the Medicare program to reflect information on comparative effectiveness would almost certainly require legislation.
- If the corresponding changes in incentives were made, generating additional information about comparative effectiveness seems likely to reduce health care spending over time—potentially to a significant degree. The precise impact, however, is difficult to predict. Given the time necessary to conduct the research, to alter incentives in a manner reflecting the results, and to affect behavior through changes in information and incentives, significant cost savings would probably take a decade or more to materialize.

Background on Past and Current Efforts

In weighing the options for expanding or reorganizing federal efforts on comparative effectiveness, it is useful to define what that term means and to consider the arguments for a federal role. Reviewing past and current efforts—by private and public organizations in the United States and by other countries—also sheds light on several issues and challenges likely to arise in any future U.S. efforts.

What Is Comparative Effectiveness?

As applied in the health care sector, an analysis of comparative effectiveness is simply a comparison of the impact of different options that are available for treating a given medical condition for a particular set of patients. Such studies may compare similar treatments, such as competing drugs, or they may analyze very different approaches, such as surgery and drug therapy. The analysis may focus only on the relative medical benefits and risks of each option, or it may go on to weigh both the costs and the benefits of those options. In some cases, a given treatment may be found more effective for all types of patients, but more commonly a key issue is determining which specific types would benefit most from it.

Although some information about the effectiveness of new drugs, medical devices, or procedures is often available, rigorous comparisons of different treatment options are less common. Drugs and devices must be certified as safe and effective by the Food and Drug Administration (FDA) before they can be marketed in the United States, but with certain exceptions the regulatory process for approving new drugs and devices does not evaluate them relative to

alternatives. Furthermore, physicians commonly prescribe drugs for “off-label” uses—that is, for treating patients or conditions that have not been certified by the FDA.³ Medical procedures, which account for an even larger share of total health care spending, can achieve widespread use without a systematic review. It may seem reasonable to assume that the benefits of a drug, device, or procedure will be similar for related conditions or a broader group of patients, and in many cases that may be true. Without hard evidence, however, decisions about what treatments to recommend often depend on the individual experience and judgment of physicians.

A recent example of a comparative effectiveness study indicates that careful analysis can sometimes disprove widely held assumptions about the relative merits of different treatments. The study, which involved patients who had stable coronary artery disease, compared the effects of two treatments: an angioplasty with a metal stent combined with a drug regimen versus a drug regimen alone.⁴ Patients were randomly assigned to receive the two treatments, and although the study found that patients treated with angioplasty and a stent had better blood flow and fewer symptoms of heart problems subsequently, the differences declined over time.⁵ More surprisingly, it found no differences between the two groups in survival rates or the occurrence of heart attacks over a five-year period.

That study examined only the comparative medical benefits of two treatments, but the term comparative effectiveness can also encompass studies that seek to determine which treatment is most cost-effective. Such studies seek to weigh any additional medical benefits of a more expensive treatment against their added costs. The benefits of different treatments are summarized as an increase in life expectancy or, more commonly, as an increase in quality-adjusted life years (QALYs) to account for effects on morbidity as well as mortality. By convention, cost-effectiveness analyses report results as the cost per QALY gained—so a lower number indicates a more cost-effective service. Related terms include cost-benefit analysis, technology assessment, and evidence-based medicine, although the latter concepts may not take costs into account.

-
3. For drug manufacturers, the costs of conducting additional trials to demonstrate safety and efficacy for a broader set of patients or conditions may outweigh the benefits from the increased sales that would result.
 4. In an angioplasty, a small balloon is surgically inserted into a clogged artery and then inflated to expand the opening; a stent—a small wire mesh tube—is added in an effort to keep the artery open.
 5. William E. Boden and others, “Optimal Medical Therapy With or Without PCI for Stable Coronary Disease.” *New England Journal of Medicine*, vol. 356, no. 15 (March 2007), pp.1503–1516. Other studies have found that angioplasty with a stent has clear medical benefits for patients who are undergoing a heart attack, reinforcing the point that results may differ among different types of patients.

Research in the Private Sector

In the United States, the private sector produces some assessments and comparisons of different treatments. One prominent source is the Technology Evaluation Center that is part of the Blue Cross Blue Shield Association. Its analyses are based on systematic reviews of the available literature, often relying on the results of clinical trials. The center produces about 20 to 25 new assessments of drugs, devices, and other technologies each year; the analyses consider clinical effectiveness but not cost-effectiveness. For-profit private-sector firms that specialize in technology assessments represent another source of information; the ECRI Institute and Hayes, Inc., are two of the larger firms providing that type of analysis. They evaluate medical and surgical procedures, drugs, and devices in return for a fee or on a subscription basis.

Organizations that are similar but operate as nonprofit entities—sometimes affiliated with academic or medical centers—include the Center for Medical Technology Policy and the Tufts-New England Medical Center’s Cost-Effectiveness Analysis Registry (which provides an extensive list of the cost-effectiveness ratios that are available from published studies). In addition, some private health plans (most commonly, larger ones) use claims data for their enrollees to conduct their own analyses of comparative effectiveness.

Notwithstanding those current efforts, the private sector will probably not produce as much research on comparative effectiveness as society would value. The knowledge created by such studies is costly to produce—but once it is produced, it can be disseminated at essentially no additional cost, and limiting that dissemination may be difficult. As a result, private insurers and other organizations conducting research on comparative effectiveness might stand to capture only a portion of the resulting benefits and therefore would not invest as much in such research as they would if they took into account the benefits to all parties. In such a situation, economists have long recognized the need for government efforts to increase the supply of research to the socially optimal level.

Another reason for the limited availability of information on comparative effectiveness is that public-sector health insurance programs—which collectively account for about half of all health care spending—have not sought to make extensive use of it. In particular, the Medicare program has not taken costs into account in determining what services are covered and has made only limited use of comparative effectiveness data. It stands to reason that the limited demand for such research has reduced the supply correspondingly. Conversely, increasing the amount of credible and objective research that was available could facilitate moving Medicare toward what former program administrator Mark McClellan has called a “fee-for-value” system rather than a fee-for-service one.

Past and Current Federal Efforts to Assess Medical Treatments

In the United States, the federal government has a rather long but somewhat checkered history of involvement in comparative effectiveness research and related efforts. Federal involvement in assessing the effectiveness of new medical technology dates at least to the late 1970s and the short-lived National Center for Health Care Technology. Established in 1978 as part of the Department of Health, Education, and Welfare, it was given a broad mandate to conduct and promote research on health care technology, and it included an advisory board appointed by the Secretary to assist in setting research priorities. The center ceased operations in 1981, however, reflecting both changes in priorities for the new Administration and the Congress as well as opposition from some provider groups.⁶ In that same period, the Office of Technology Assessment (OTA) was created as an advisory agency to the Congress covering a broad set of issues, including but not limited to health care. Over the years, it studied a variety of health care topics, including the costs and benefits of screening tests for several diseases. For a variety of reasons, however—which apparently had little to do with its health care studies—OTA was eliminated in 1995.

More recently, the Agency for Health Care Research and Quality (AHRQ) has been the primary federal agency supporting research on the effectiveness of medical treatments and their comparative benefits and costs. Established in 1989 as the Agency for Health Care Research and Policy, AHRQ is an arm of the Department of Health and Human Services (HHS). It currently has a staff of about 300 and an annual budget of about \$300 million, which primarily funds research grants to and contracts with universities and other research organizations covering a wide range of topics in health services.

AHRQ has undertaken a number of initiatives related to comparative effectiveness. One such step—in collaboration with the American Medical Association and America’s Health Insurance Plans, a coalition of insurance companies—has been the creation of a national clearinghouse for treatment guidelines, which are designed to summarize the available medical evidence on appropriate treatments for various conditions. AHRQ has also established about a dozen evidence-based practice centers around the country, generally with an affiliation to a university; those centers analyze and synthesize existing evidence about treatments and technologies. Some studies sponsored by AHRQ have examined only the relative clinical benefits of different treatments, while others have also analyzed their cost-effectiveness. Research on comparative effectiveness accounts for only a portion of AHRQ’s budget, however.

As with other agencies examining the effectiveness of medical treatments or evaluating medical technologies, support for AHRQ has varied over time. In the mid-1990s, controversies arose in connection with a panel that was seeking to review evidence and formulate guidelines about the treatment of back pain, and

6. See Seymour Perry, “The Brief Life of the National Center for Health Care Technology,” *New England Journal of Medicine*, vol. 307, no. 17 (October 21, 1982), pp.1095–1100.

partly as a result, the agency faced the prospect of elimination. Ultimately, the agency was retained, but its funding was reduced from prior levels. Since then, its overall budget has generally been maintained (at least in nominal terms) or increased. Most recently, section 1013 of the Medicare Modernization Act of 2003 authorized AHRQ to spend up to \$50 million in 2004 (and additional amounts in future years) to conduct and support research with a focus on “outcomes, comparative clinical effectiveness, and appropriateness of health care items and services (including prescription drugs)” for Medicare and Medicaid enrollees. Actual funding for that initiative has been \$15 million per year.

Other federal agencies also engage in various activities related to comparative effectiveness research. The National Institutes of Health (NIH)—also part of HHS—is the largest federal sponsor of clinical research, primarily in the form of clinical trials. Although comparative effectiveness is not a focus of its research, over the years a number of trials have been sponsored that compare treatments head to head. The Department of Veterans Affairs also has a substantial research program that reviews evidence from the clinical records of its patients, focusing particularly on the clinical effectiveness of treatments. The department also sponsors evidence reviews through a technology assessment program.

The Centers for Medicare and Medicaid Services (CMS) has also sponsored some research on comparative effectiveness. When making decisions about what services are covered by Medicare, CMS generally considers only whether devices and procedures are effective. It has sponsored some studies comparing the effectiveness of different treatments but has done so largely to determine whether to establish separate payment rates for similar treatments. For example, CMS is currently cosponsoring a trial with NIH that may eventually compare the effects of daily dialysis for kidney patients with the conventional treatment of dialysis three times per week. If daily dialysis proves more effective for certain patients, CMS could modify its payment policy to cover the additional costs of more frequent treatment.

Comparative Effectiveness in Other Countries

Other developed countries also face challenges financing health care costs and have taken various steps to assess the comparative effectiveness of treatments. Many of those countries establish overall budgets for their national health systems and use comparative effectiveness analysis to help determine which treatments and procedures will be covered or how they will be reimbursed. Perhaps the best known example is the National Institute for Health and Clinical Excellence (NICE), which was established in 1999 as part of the United Kingdom’s national health service. It provides guidance on the use of new and existing medicines, procedures, and treatments and on appropriate treatments for specific diseases. With a staff of about 200 and an annual budget of about 30 million pounds (roughly \$60 million), NICE does not fund new clinical trials or other forms of primary data collection but, rather, bases its determinations on systematic reviews of existing research.

Other countries such as Australia, Canada, France, and Germany have similar review processes. Discussions have sometimes focused on those countries' procedures for reviewing prescription drugs, but all of them have systems in place to evaluate medical and surgical treatments and technologies as well. It is therefore worth noting that, for all the attention that prescription drugs receive, they currently account for less than 15 percent of total U.S. health spending. Therefore, if additional research in the United States on comparative effectiveness focused only on medications, the impact would probably be much smaller than if that research encompassed the whole spectrum of medical care.

Options for Organizing and Funding Federal Research Efforts

The approach that is taken for organizing and funding any increased federal efforts to support research on comparative effectiveness will play an important role in determining their impact. Many of the options that have been proposed seek to coordinate and centralize existing activities through one entity—which would tend to give any conclusions it reached more weight—but there might also be value in developing several competing sources of information about comparative effectiveness.

Options that have been put forward for organizing federal research on comparative effectiveness include the following (each of which could have many variants):⁷

- Expanding the role of an existing agency that already conducts or oversees research on health services generally—and comparative effectiveness specifically—such as AHRQ or NIH.
- Creating or “spinning off” a new agency, either within the Department of Health and Human Services or as an independent body that is part of either the executive or the legislative branch. The Federal Trade Commission and the Medicare Payment Advisory Commission (MedPAC) are potential models for such an option.
- Augmenting an existing quasi-governmental organization, such as the Institute of Medicine or the National Research Council. Such entities are often Congressionally chartered, but they are not subject to regular governmental oversight. Even so, the Institute of Medicine receives most of its funding from government agencies, which is provided to finance specific studies that have been requested.

7. For a discussion of this issue, see Gail R. Wilensky, “Developing a Center for Comparative Effectiveness Information,” *Health Affairs*, Web Exclusive (November 7, 2006), pp. w572–w585.

- Establishing a new public-private partnership to oversee and direct research. That option could be structured in various ways, but one such approach would be to set up a federally funded research and development center (FFRDC). FFRDCs are not-for-profit organizations that can accept some private funding but which get most of their funding from a federal agency that provides oversight and monitoring.

Regardless of how those efforts were organized, several potential mechanisms could be used to fund them (either individually or in combination). Federal spending could be authorized and appropriated annually, as with other discretionary programs. Alternatively, funding could be drawn from Medicare's hospital insurance trust fund (which is financed primarily by payroll taxes) or specified as a percentage of mandatory federal outlays on health insurance programs. Instead of or in addition to using existing sources of revenues, another set of funding options would require direct contributions from the health sector. For example, a new tax on health insurance premiums or other payments within the health sector could be established, with the resulting revenues dedicated to research on comparative effectiveness.

A comparison of those organizational and funding options for a new or expanded entity indicates that trade-offs may arise between the entity's independence and ability to reach controversial conclusions, on the one hand, and its accountability and responsiveness to policymakers and to other interested parties, on the other. For example, funding through appropriations would allow lawmakers to assess the new entity's contributions and accomplishments and to balance spending on those efforts against other federal priorities on an annual basis. But at the same time, some observers have raised concerns that annual appropriations would leave a new entity vulnerable to outside political pressure and thus reluctant to undertake controversial studies or to reach conclusions that might generate opposition from affected groups. In that view, the elimination of agencies engaged in such research that were funded by annual appropriations—or in the case of AHRQ, the threat of elimination—suggests the need for a different arrangement.

Alternatively, housing the new activities in an organization that was at “arm's length” from the federal government, and establishing automatic or dedicated funding mechanisms, would give the new entity greater autonomy. To be sure, lawmakers could change any funding formula that had been established, mitigating the insulation from outside pressure. For example, payment rates to doctors, hospitals, private health plans, and other providers under Medicare have been adjusted frequently in response to concerns about their levels, even though such payments are not subject to annual appropriations. Conversely, to the extent that automatic or dedicated funding mechanisms did limit the influence of outside pressure, they also would raise questions about how the entity set its priorities and allocated resources—and how it would be held accountable for those decisions.

Under any option, a governing council or advisory board could be established to serve several functions: providing guidance to the entity and establishing priorities for its research projects, creating an independent process for reviewing and approving the findings that resulted from that research, and serving as a channel for interested parties to participate in its deliberations. For example, the council or board could include representatives of major federal health programs, private insurers, health care providers, and drug and device manufacturers—as well as members of the general public and disinterested policy experts. Alternatively or in addition, a regular process could be established for getting input from interested parties. The types of participants on any council or board and the manner in which members were chosen and replaced would have to be determined carefully to avoid giving one perspective undue influence. At the same time, trade-offs could arise between the extent to which a broad range of views and interests were represented and the ability of the council or board to make timely decisions or to reach consensus on contentious issues.

Another organizational issue is whether to establish a single or highly centralized entity or, instead, to design a more loosely coordinated system encompassing several distinct centers to produce independent analyses. Many of the options that have been proposed seek to centralize research activities through one entity—partly to address concerns about the lack of coordination among current U.S. efforts. An advantage of that centralized approach is that it would tend to give more weight to any conclusions reached. At the same time, that potential for having a greater impact could also lead the organization to adopt findings that were watered down to reach consensus. An alternative, more decentralized approach could give individual research centers more latitude and foster competing perspectives. However, a more pluralistic approach could involve some redundant efforts and, if it yielded any conflicting findings, would leave users with the task of reconciling the results.

An additional consideration that arises—particularly if a new entity is created—involves start-up costs and other implementation challenges. If funds were directed through an existing federal agency, some ongoing costs for additional staffing would be incurred, but the support infrastructure would largely exist already. By contrast, establishing a new agency or public-private partnership could require a greater effort before research activities could commence. At the same time, a quasi-governmental organization or public-private partnership could have more flexibility to develop and maintain its staff than a new or existing federal agency would. Creating a new source of revenues (such as a tax on health insurance premiums) to help fund research efforts on comparative effectiveness would also involve time and administrative costs.

Issues and Challenges in Comparing the Effectiveness of Treatments

The appropriate organizational form for any new or expanded federal entity, along with the mechanism and level of funding, depends in large part on what activities it would carry out. For example, analyzing existing data would require a different set of skills—and would cost less—than overseeing new clinical trials that compared different treatments. Whatever approach was taken, communicating the results of the analyses to doctors, patients, and health insurers in ways that each audience found useful would probably be an important function.

Methods of Research

The approach that would probably be easiest to implement would be to review and summarize the results of existing studies. For example, even though existing studies may only compare a single treatment to a placebo, the results of several studies could sometimes be combined to measure treatments against one another. Such activities would be comparable to some of the work that AHRQ is already undertaking and to some current efforts based at universities or other public and private research centers.

In pursuing that approach, one challenge is that some analyses have indicated that clinical trials sponsored by interested parties—which are often the only source of such data—are more likely than independent studies to find favorable results.⁸ Even without that problem, another potential limitation is that existing information may not be sufficient to reach definitive conclusions, either because studies use different methodologies or analyze different populations of patients, or simply because they yield conflicting findings. For example, there are a range of independent studies available that examine different screening techniques for colorectal cancer, each of which provides an estimate of the cost per enrollee for each increase in QALYs. But a recent review of those studies conducted by MedPAC suggests that reaching a firm conclusion about which approach is most effective or most cost-effective would be difficult because their results vary considerably.⁹

-
8. See Justin E. Bekelman, Yan Li, and Cary P. Gross, “Scope and Impact of Financial Conflicts of Interest in Biomedical Research: A Systematic Review,” *Journal of the American Medical Association*, vol. 289, no. 4 (January 22/29, 2003), pp. 454–465; Stephan Heres and others, “Why Olanzapine Beats Risperidone, Risperidone Beats Quetiapine, and Quetiapine Beats Olanzapine: An Exploratory Analysis of Head-to-Head Comparison Studies of Second-Generation Antipsychotics,” *American Journal of Psychiatry*, vol. 163, no. 2 (February 2006), pp. 185–194; and Jeffrey Peppercorn and others, “Association Between Pharmaceutical Involvement and Outcomes in Breast Cancer Clinical Trials,” *Cancer*, vol. 109, no. 7 (April 2007), pp. 1239–1246.
 9. Medicare Payment Advisory Commission, *Report to the Congress: Increasing the Value of Medicare* (June 2006), pp. 232–233.

In other cases, though, the existing evidence may permit more clear-cut determinations. Britain's NICE, for example, has been able to analyze many different treatments on the basis of their cost-effectiveness and to develop an extensive set of clinical guidelines and technology assessments relying solely on systematic reviews of available studies. It is also worth noting that "inconclusive" studies or comparisons may still be valuable, in that they could simply indicate that different therapies have equivalent benefits or are equally cost-effective. If, instead, inconclusive results reflect substantial differences in the studies' findings or a lack of sufficient analysis using rigorous methodologies, then generating new evidence may be necessary.

A somewhat more challenging approach than reviewing existing studies would be to fund new analyses comparing medical treatments using existing data sources, such as health insurance claims records. An advantage of that approach is that it could provide new information to help resolve uncertainties about treatments at relatively low cost. A central difficulty in such studies, however, is accounting for differences in patients' health status that play a role in determining which treatment they get—which can make simple comparisons misleading. Patients with more severe heart disease, for example, are more likely to receive invasive and expensive surgical procedures such as angioplasty or a bypass operation. But the greater severity of their condition may also make them more likely to have a subsequent heart attack and more likely to die, so a comparison to patients receiving less aggressive treatments—who are not as sick, on average, to begin with—could understate the benefits of more aggressive treatments. To address such problems, researchers might be able to exploit geographic differences in treatment patterns to compare the effects of different treatments on comparable types of patients. Expanded use of electronic health records could also facilitate more sophisticated statistical analyses, assuming that issues regarding access to and privacy of those records could be addressed.

The method of research that would probably yield the most definitive results is one employing randomized controlled trials comparing treatments head to head, but that approach would also be the most expensive and would take the longest to conduct. The main advantage of random assignment is that it ensures that any differences in outcomes reflect true differences among treatments and not confounding differences among patients. But detecting differences that are statistically significant—that is, unlikely to have arisen simply by chance—can require a substantial number of patients to participate, and they must generally be followed for several years. Ethical issues can also arise if one set of participants is assigned a treatment that is generally considered less effective, although such concerns are less likely to arise when significant uncertainty exists in the medical community about the relative benefits of different treatments. Because the number of trials that could feasibly be conducted at any given time is limited, significantly expanding comparative effectiveness research would therefore be likely to require a combination of randomized trials and other research methods.

Scope of Analysis and Dissemination of Results

The results of clinical trials and other comparisons of treatment options will address most directly the relative medical benefits of those options, but an important question is whether federal research on comparative effectiveness would also seek to assess the cost-effectiveness of treatments. There are arguments both for and against doing so.

An argument against having the federal entity assess which treatments are most cost-effective is that doing so may be unnecessary. If that entity generated detailed data about the treatments studied (including any ripple effects on other types of care, such as hospital readmissions), health plans and others could use relevant prices to calculate cost-effectiveness ratios. Furthermore, the evidence that some areas of the country spend substantially more on health care but do not have better health as a result suggests that improvements in the efficiency of health care delivery could be obtained without having to address trade-offs between costs and benefits—with the first step being to reduce the use of those treatments that do not provide more medical benefits than alternative, less expensive therapies.

Ultimately, however, achieving the greatest possible gains in the efficiency of the health sector would require assessing both the benefits and costs of different treatments to see whether the added benefits of more-expensive options were worth their added costs. Having the new or expanded entity analyze cost-effectiveness would also lend more legitimacy to that approach and would promote a consistent and transparent methodology for such calculations. Those considerations argue for having the entity assess cost-effectiveness.

Another question is whether assessments would be limited to procedures and treatments or would also seek to evaluate the performance of individual doctors. In particular, the data from medical records that are used to compare the effectiveness of different treatments for a given type of patient could also be used to analyze the quality with which doctors provided each treatment. The potential gains from such analysis include identifying doctors who deliver high-quality care and encouraging doctors who are not performing as well to improve—and doing both on the basis of objective evidence. At the same time, concerns could arise that evaluating doctors would detract from the focus on identifying effective procedures. Further, controlling for differences among patients that could affect individual doctors' ratings could be even more challenging than controlling for differences among patients when comparing treatments. Although such an approach could have a larger impact on health care than examining treatments alone, it could also be highly controversial.¹⁰

10. CMS has taken some initial steps toward assessing the quality of care that individual doctors provide. The Tax Relief and Health Care Act of 2006 allows for modest bonus payments under Medicare to doctors who elect to report information to CMS on certain measures of the care they provide in 2007. Although CMS will be able to provide feedback to doctors on how their performance compares to their peers', the payments

Whichever method of research was used to generate new findings, achieving consensus about their implications could prove challenging. For example, although there may be substantial agreement within the scientific community about the relative benefits of avoiding different adverse outcomes—such as degrees of disability and risks and side-effects of surgery—converting those differences into the metric of quality-adjusted life years may nevertheless raise concerns among patients and other interested parties. Similarly, deciding how broadly or narrowly any findings apply would be an important consideration, because some treatments may be more effective for certain subgroups of patients than for an average patient. That consideration would also affect the design of the studies and the comparisons that would be undertaken. Finally, determining which treatment was most cost-effective for a given population would involve placing a dollar value on an additional year of life, which has in the past generated public controversy—even though researchers have developed estimates of that value reflecting choices that individuals are observed to make.

Communicating the results of research to doctors, patients, and health insurers would probably be an important activity for any new or augmented entity focused on comparative effectiveness. Providing information to both technical and nontechnical audiences that was useful and accurate would be challenging. A particular difficulty might be conveying the degree of uncertainty surrounding conclusions.

Potential Implications for Health Care Spending

To affect medical treatment and reduce health care spending, the results of comparative effectiveness analyses would ultimately have to change the behavior of doctors and patients. For any large-scale changes to occur, the new or expanded entity would have to generate new findings for a substantial number of medical conditions—which would take many years. To affect behavior, those findings would then probably have to be incorporated into the incentives for providers and patients, a process of adjustment that might also take time.

Medicare is effectively precluded from taking cost into account when making decisions about coverage and would probably need new legal authority to adjust payments to providers or cost-sharing requirements for enrollees to encourage the use of more cost-effective care. For their part, private insurers might not face legal barriers to limiting coverage of treatments that were shown to be less effective but still might be reluctant to do so if Medicare continued to cover them. Beyond the analyses themselves, then, many steps would need to be taken before

doctors receive do not depend on that performance, and the measures that have been chosen cover areas of substantial consensus in the medical community about appropriate treatment protocols (for example, prescribing beta blockers to patients who have had a heart attack).

spending on comparative effectiveness translated into savings for federal programs and the health care system.

Potential for Savings on Health Care

Predicting the impact that research on comparative effectiveness could have on health care spending is difficult because it is hard to know what that research will show. In some cases, the research could provide clearer evidence than exists today that the benefits of a treatment (such as a screening examination) outweigh the costs—in which case spending on such treatments could increase.

As a general rule, however, the fee-for-service reimbursement system by which health care is currently financed—especially in Medicare—typically provides financial incentives for doctors and hospitals to adopt new treatments and procedures broadly even if hard evidence about their effectiveness is not available. For their part, insured individuals generally face only a portion of the costs of their care and, consequently, have only limited financial incentives to seek a lower-cost treatment. Private health insurers have incentives to limit the use of ineffective care but are currently constrained by a lack of information, by the turnover of enrollees when they change insurance coverage, and by public concerns about overly aggressive management (as was evident in the recent “backlash” against managed care plans). Over the long term, therefore, additional objective information about the relative costs and benefits of treatments—if adopted by insurers and accepted by doctors and patients—seems more likely to reduce total health care spending than to raise it.

Getting to the point where additional research on comparative effectiveness could have a noticeable impact on health spending would itself take several years. In addition to the time required to get the new activities under way, a lag would exist before results were generated—particularly if they depended upon new clinical trials. Initially, the available results would probably address a relatively small number of medical treatments and procedures; additional time would have to elapse before a substantial body of results was amassed. And in areas of medicine that involve significant levels of spending, several studies could be needed before a consensus emerged about the appropriate conclusions to be drawn—even if those studies did not generate conflicting results. For all of those reasons, it would probably be a decade or more before new research on comparative effectiveness had the potential to reduce health care spending in a significant way.

Possible Responses Under Private and Public Insurance Plans

To affect medical treatment and reduce health care spending, the results of comparative effectiveness analyses would ultimately have to change the behavior of doctors and patients—that is, to get them to use fewer services or less intensive and less expensive services than are currently projected. Bringing about those changes would probably require action by public and private insurers to incorporate comparative effectiveness information into their coverage and payment policies in order to affect the incentives facing doctors and patients.

Although private insurers could choose not to cover drugs, devices, or procedures that were found to be less effective or less cost-effective, the insurers would have a number of additional options as well. They could simply provide more information to patients and doctors or make that information public, which could improve compliance with treatment guidelines. For example, the use of medicines known as beta blockers, which is recommended following a heart attack to prevent recurrence, has grown substantially in recent years—apparently as a result of requirements for health plans to report the share of patients who receive prescriptions for them.¹¹ Alternatively, insurers could require enrollees to pay some or all of the additional costs of more-expensive treatments that were shown to be less effective or less cost-effective (in which case enrollees would have to decide whether the added benefits were worth the added costs). Or insurers could adjust payments to doctors and hospitals to encourage the use of more-effective care.

The steps that private insurers took could both affect public spending and be affected by public programs' responses to additional information about comparative effectiveness. To the extent that changes instituted by private insurers affected doctors' methods, there could be spillover benefits for public programs—because physicians typically serve patients of both types of programs and tend to use the same general approach to care. However, private insurers might be reluctant to pursue such approaches aggressively if public insurance programs were not adopting similar methods. In addition, private insurers might be slow to cover treatments or screening exams that took a long time to generate savings in other health costs, either because of turnover in their membership or because of questions about the benefits of providing coverage for routine services. (A preventive service could be cost-effective overall but the additional savings to the insurer that result from covering it could still be less than the costs of providing that coverage to all enrollees.)

To reduce spending under Medicare on the basis of comparative effectiveness research would very likely require additional legislative authority both to allow relative benefits and costs to be considered and to modify the financial incentives in that program. Under current law, Medicare does not appear to have the authority to take costs into account when making decisions about what treatments are covered; regulations have been proposed in the past that would have used

11. Since 1996, the National Committee for Quality Assurance (NCQA), a not-for-profit organization that provides information about health care quality, has required private health care plans to report that information. The average share increased from 63 percent in 1996 to 95 percent in 2005, and as a result, NCQA has now adopted a more stringent measure (which tracks actual use of those drugs). See National Committee for Quality Assurance, "New HEDIS® Measures Track Childhood Lead Screening, COPD Management; Retirement of Beta-Blocker Measure Marks Major Accomplishment in Cardiac Care" (news release, February 21, 2007), available at web.ncqa.org/tabid/254/Default.aspx.

costs as a factor, but those proposals generated opposition and were ultimately withdrawn.¹² As a result, Medicare will generally cover any treatment or procedure that has medical benefits, regardless of its cost or its effectiveness relative to alternative therapies. Recently, Medicare officials developed an initiative that provides provisional coverage for new treatments that have uncertain medical benefits—but also requires the resulting evidence about their effects to be analyzed so that a more informed final decision on coverage can be made using those data. That approach, however, does not involve comparing different treatments to see which is more effective, nor does it take the costs of treatments into account.

Medicare currently has somewhat more flexibility regarding the payments it makes for covered services, which can take comparative medical benefits (but not costs) into account on a limited basis. For example, in order for a hospital to receive an additional payment for using a new device (known as a “pass-through” payment), the device must be shown to provide a substantial clinical improvement for Medicare beneficiaries compared with currently available treatments. (Over time, payments to hospitals for new technologies are incorporated into Medicare’s prospective payment rates.) In addition, Medicare has adopted a “least costly alternative” payment policy for certain types of items, under which it will not cover the additional cost of a more expensive product if a clinically comparable one is available that costs less. That policy has been applied to payment for durable medical equipment and to certain comparable drugs, but wider application to services such as surgeries or other treatments and procedures would probably require additional authority.¹³

If the necessary changes in law were made, Medicare could use information about comparative effectiveness to promote higher-value care. For example, Medicare could tie its payments to providers to the cost of the most effective or most efficient treatment. If that payment was less than the cost of providing a more expensive service, then doctors and hospitals would probably elect not to provide it—so the change in Medicare’s payment policy would have the same practical effect as a coverage decision. Alternatively, enrollees could be required to pay for the additional costs of less effective procedures (although the impact on patients’ incentives and their use of care would depend on whether and to what extent they had supplemental insurance coverage that paid some or all of Medicare’s cost-sharing requirements).

12. The Medicare statute essentially requires that program to cover any items or services that are deemed “reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member.” See section 1862(a)(1)(A) of the Social Security Act.

13. For further discussion about the use of information about comparative effectiveness under Medicare, see Medicare Payment Advisory Commission, *Report to the Congress: Issues in a Modernized Medicare Program* (June 2005), pp. 180–182.

More modest steps that Medicare could take would include smaller-scale financial inducements to doctors and patients to encourage the use of cost-effective care. Doctors and hospitals could receive modest bonuses for practicing effective care or modest cuts in their payments for using less effective treatments (although the evidence to date about the effect of such pay-for-performance initiatives on health spending is somewhat mixed).¹⁴ Likewise, enrollees could be asked to pay only a portion of the additional costs of less efficient procedures. Or Medicare could simply provide information to doctors and their patients about their practices, which would create some pressure for doctors to use more-efficient approaches. Adopting more modest measures to incorporate the findings of comparative effectiveness research, however, is likely to yield smaller savings for the program. It is also worth noting that under current law, policies that could affect the use of physicians' services by Medicare enrollees will not change the program's spending unless the targets for spending (under what is known as the Sustainable Growth Rate system) are also changed; otherwise, payment rates for physicians would be adjusted automatically to keep total spending unchanged.

As for Medicaid, state officials generally determine what specific services are covered—subject to broad federal requirements—and are reimbursed by the federal government for a portion of the resulting costs using formulas specified in law. Because enrollees have low income, options for adjusting cost-sharing requirements to encourage the use of cost-effective care may be limited. Furthermore, a substantial portion of Medicaid spending pays for long-term care services such as nursing homes for elderly and disabled enrollees, which would probably not be affected by comparative effectiveness research. At the same time, most of the poor mothers and children enrolled in the program receive their care through a private health insurance plan under contract to Medicaid, so spending for them would be directly affected by any changes that private insurers made. Another portion of Medicaid spending goes to cover cost-sharing requirements and payments of premiums for enrollees who are also on Medicare, so the impact on that spending would depend largely on what the Medicare program did.

An additional issue that would arise in applying the results of comparative effectiveness studies in Medicaid is that states would ordinarily stand to capture only a portion of any savings that were generated. Federal matching rates under Medicaid currently range from 50 percent up to about 75 percent, and, by CBO's estimates, the federal government now covers 57 percent of the costs of health services provided by that program. In principle, those financing arrangements would reduce incentives for state Medicaid officials to limit coverage of less effective services—because, on net, states would see only a portion of those costs (or a portion of the savings). Some coordination between state and federal officials might therefore be required to incorporate the results of comparative effectiveness research. At the same time, many states recognize the growing fiscal

14. See Congressional Research Service, *Pay-for-Performance in Health Care*, CRS Report RL33713 (December 12, 2006).

burden posed by Medicaid costs, and several of them have already expressed interest in such research. For example, more than a dozen state Medicaid programs are involved in a project (affiliated with the Oregon Health and Sciences University) assessing evidence about the relative safety and effectiveness of competing drugs in the same class.

Conclusion

The United States is on an unsustainable fiscal path, and the primary determinant of the nation's long-term fiscal balance is health care costs. The substantial variation in health care costs per beneficiary across the nation, in a fashion unrelated to health outcomes, strongly suggests that opportunities exist to reduce those costs without impairing health—underscored by the lack of relative evidence on “what works and what doesn’t” for many health care interventions. Expanded research on comparative effectiveness, if linked to changes in incentives for providers and patients, offers a promising mechanism for reducing health care costs to a significant degree over the long term while maintaining or improving the health of Americans.